解答例

STEP I 2007、第 13 問

確率・統計 ・ 条件付き確率

> **STEP I 2007、問 13**
>
> 先に進む前に、まず問題を読んでおこう。問題は STEP Support Programme の問題データベースにある：
>
> step.maths.org/questions/07-s1-q13

この問題は、*Beyond STEP* の第 12.1 節「**正確な条件付け**」で詳しく解説されている：
https://beyondstep.keithmcnulty.org/ja/ch12/conditioning.html

鍵となる考え方

円板は 1 枚ずつ取り出されるが、4 枚の円板のどの組も等しい確率で現れるので、作業の大部分は組合せ（combinations）を数えることである。面白いのは条件のほうである。**どの**円板が最初に出たかを知ること、あるいは単にある特定の円板が取り出された**ということ**を知ることは、残る 3 枚の円板を、残りの 10 枚からの無作為な選択のままにしておく。一方、**何らかの**番号付きの円板が取り出されたことだけを知るのは、まったく種類の異なる情報であり、最後の二つの小問はその違いを示すために置かれている。

解答

取り出した 4 枚のうち番号付きの円板の枚数を N とする。袋はよく振られているので、相異なる 4 枚の円板の順序付きの並びはどれも等しい確率で現れ、したがって順序を考えない 4 枚の組もどれも等しい確率で現れる。そのような組は

$$\binom{11}{4} = 330$$

通りある。番号付きの円板をちょうど k 枚含む組は、5 枚の番号付きの円板から k 枚、6 枚の無地の円板から $4-k$ 枚を選ぶことで作られるので、

$$\mathrm{P}(N=k) = \frac{\binom{5}{k}\binom{6}{4-k}}{330}.$$

小問 (i)

4 枚すべてが番号付きであるのは $k = 4$ のときである：

$$\mathrm{P}(N = 4) = \frac{\binom{5}{4}\binom{6}{0}}{330} = \frac{5}{330} = \frac{1}{66}.$$

同じことを、順に取り出すと考えて計算すれば $\frac{5}{11} \cdot \frac{4}{10} \cdot \frac{3}{9} \cdot \frac{2}{8} = \frac{1}{66}$ となる。

小問 (ii)

円板「3」が最初に取り出されるという事象を T_1 とする。T_1 のもとでは、振られた袋には 10 枚の円板、すなわち番号付きが 4 枚、無地が 6 枚入っており、次の 3 枚はこれらからの無作為な選択である。4 枚すべてが番号付きになるのは、この 3 枚がすべて番号付きであるときに限るので、

$$\mathrm{P}(N = 4 \mid T_1) = \frac{\binom{4}{3}}{\binom{10}{3}} = \frac{4}{120} = \frac{1}{30}.$$

小問 (iii)

引き続き T_1 のもとで、残りの 10 枚から取り出す 3 枚のうちに、番号付きの円板がちょうどもう 1 枚含まれればよい：

$$\mathrm{P}(N = 2 \mid T_1) = \frac{\binom{4}{1}\binom{6}{2}}{\binom{10}{3}} = \frac{4 \times 15}{120} = \frac{1}{2}.$$

小問 (iv)

今度は、円板「3」がどこかの時点で取り出されるという事象を T とする。

> **考えてみよう** 円板「3」が最初に出たか、それ以外の回に出たかは問題になるだろうか？

問題にならない。T のもとでは、取り出された 4 枚は円板「3」とほかの 3 枚からなり、対称性（symmetry）により、ほかの 10 枚から選ばれる 3 枚の組はどれも等しい確率で現れる。これはまさに小問 (iii) の状況であり、そこでは取り出す順序はまったく関係なかった。したがって答えはやはり $\frac{1}{2}$ である。

直接確かめてみよう。$\mathrm{P}(T) = \binom{10}{3}/\binom{11}{4} = \frac{120}{330} = \frac{4}{11}$ であり、円板「3」、ほかの番号付きの円板 1 枚、無地の円板 2 枚を含む組は $\binom{4}{1}\binom{6}{2} = 60$ 通りあるので、

$$\mathrm{P}(N = 2 \mid T) = \frac{60/330}{120/330} = \frac{1}{2}.$$

小問 (v)

最初の円板が番号付きであるという事象を F とする。最初の円板にどの番号が書かれているかによって F を分ける。F は五つの事象「円板 j が最初に取り出される」($j = 1, \ldots, 5$) の互いに素な和（disjoint union）であり、それぞれの確率は $\frac{1}{11}$ である。小問 (iii) の議論は番号が 3 であることを使っていなかったので、各 j に対して

$$\mathrm{P}(N = 2 \mid \text{円板 } j \text{ が最初}) = \tfrac{1}{2}.$$

よって、全確率の法則（law of total probability）により、

$$\mathrm{P}(N = 2 \mid F) = \sum_{j=1}^{5} \frac{1/11}{5/11} \cdot \frac{1}{2} = \frac{1}{2}.$$

（直接考えれば、F のもとでほかの 3 枚は番号付き 4 枚と無地 6 枚からの無作為な 3 枚であり、これはふたたび小問 (iii) である。）

小問 (vi)

番号付きの円板が少なくとも 1 枚取り出されるという事象を A とする。すなわち A は事象 $N \geq 1$ である。これはこれまでと同じ種類の情報**ではない**。1 枚の円板を特定して残りを無作為のままにするわけではないからである。そこで条件付き確率（conditional probability）の定義を用いる。$N = 2$ ならば A が成り立つので、

$$\mathrm{P}(N = 2 \mid A) = \frac{\mathrm{P}(N = 2)}{\mathrm{P}(N \geq 1)}.$$

考えてみよう なぜ単に「番号付きの円板が 1 枚わかっているので、ほかの 3 枚は残りの 10 枚から無作為に選ばれる」と言えないのだろうか？

どの円板がわかっている 1 枚なのかを知らないからである。番号付きの円板を 2 枚含む組も、4 枚含む組も、同じように A を満たす。脇に取り分けておける特定の 1 枚は存在しない。したがってこの条件は N そのものに応じて各組の重みを付け直すことになり、数え上げなければならない。

$$\mathrm{P}(N = 2) = \frac{\binom{5}{2}\binom{6}{2}}{330} = \frac{10 \times 15}{330} = \frac{150}{330}, \qquad \mathrm{P}(N = 0) = \frac{\binom{6}{4}}{330} = \frac{15}{330},$$

であるから、$\mathrm{P}(N \geq 1) = \frac{315}{330}$ となり、

$$\mathrm{P}(N = 2 \mid A) = \frac{150}{315} = \frac{10}{21}.$$

検算

N の分布（distribution）は合計が合うはずである。$k = 0, \ldots, 4$ に対する個数 $\binom{5}{k}\binom{6}{4-k}$ は $15, 100, 150, 60, 5$ であり、合計は 330 となる。小問 (ii) と (iii) では、追加の番号付きの円板の枚数の条件付き分布の個数は 120 通りのうち $20, 60, 36, 4$ であり、こちらも正しく合計が合う。さらに、$11 \times 10 \times 9 \times 8 = 7920$ 通りの順序付きの取り出し方をすべて書き出し、正確な分数で計算することで、六つの答え（$\frac{1}{66}$、$\frac{1}{30}$、$\frac{1}{2}$、$\frac{1}{2}$、$\frac{1}{2}$、$\frac{10}{21}$）をすべて確かめた。$\frac{10}{21} \approx 0.476$ が $\frac{1}{2}$ より少し小さいのは当然であることに注意しよう。$N \geq 1$ という条件は $N = 0$ の場合を取り除くだけなので、特定の円板がわかっている場合に比べて $N = 1$ に多くの重みが残るのである。

よくある間違い

- 小問 (iv) と (vi) を同じ種類の条件として扱ってしまう。「円板 3 が取り出される」は円板を名指ししているが、「番号付きの円板が取り出される」はそうではなく、答えも異なる。
- 小問 (vi) で、条件が $N \geq 1$ であることを忘れ、代わりに $\mathrm{P}(N = 1)$ や $\frac{5}{11}$ で割ってしまう。
- 分子では順序付きの取り出し方を数え、分母では順序を考えない組を数えてしまう。どちらか一方を選び、それを貫くこと。

STEP のその先へ

小問 (v) と (vi) の対比は、有名な「男の子か女の子か」のパラドックスと同じ落とし穴である。**特定の**子どもが男の子だと教えられた場合と、**少なくとも一人**が男の子だと教えられた場合とでは、答えが異なる。大学では、事象による条件付けを確率測度（probability measure）の制限と正規化の取り直しとして扱うことで、このことが厳密に定式化される。また、ここでの個数は超幾何分布（hypergeometric distribution）にほかならない。Grimmett と Welsh の *Probability: An Introduction*（確率論入門）、および Blitzstein と Hwang の *Introduction to Probability*（確率論入門）は、いずれもこうしたパラドックスを丁寧に論じている。

Beyond STEP 解答集：https://beyondstep.keithmcnulty.org/ja/。解答本文は CC0 1.0 である。問題そのものは © Cambridge University Press & Assessment に帰属し、ここには掲載していない。英語版から機械翻訳したものであり、まだ校閲を経ていない。